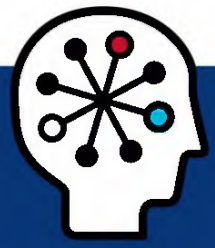
HP Power Capping and HP Dynamic Power Capping for ProLiant servers



Abstract 3

Introduction 3

Basics of server power control 3
- Processor P-states 4
- Clock throttling 5

How power capping functions 6
- Maintaining power consumption below the cap 6
- Minimum and maximum power consumption for a server 6

Differences between HP Dynamic Power Capping and HP Power Capping 7
- Power provisioning and Dynamic Power Capping 7
- Support for Power Capping in ProLiant servers 8

Group power capping for servers through Insight Power Manager 8

Enclosure Dynamic Power Capping 9
- Elements of an enclosure power cap 9
- Enclosure Dynamic Power Capping operation 10
- Active power reallocation 11
- Enclosure Dynamic Power Capping in mixed blade environments 11
- Opting out servers 11

Setting power caps for servers 12
- Setting a power cap for a single server 12
- Setting a power cap for a group of servers 14
- Setting an BladeSystem enclosure power cap 15
- Setting a power cap for a group of enclosures 16

Using power capping in data center provisioning 17
- Choosing effective power caps 17
- Power capping to peak power consumption 19
- Power capping to average power consumption 21

Using Enclosure Dynamic Power Capping in power provisioning 21

Additional uses for power capping 22
- Power capping for emergency management 22
- Time-of-day power capping 23

Subtleties of power capping 24
- Avoiding power capping conflicts within groups 24
- Powering-up groups of servers when using Dynamic Power Capping 24
- Setting low or unattainable power caps on servers 24
- Peak power reporting and Dynamic Power Capping 25

Downloaded from www.Manualslib.com manuals search engine

Using HP Pawer Regulatar in canjunctian with pawer capping 25
Pawer capping and CPU utilizatian 25
Summary 26
Far mare infarmatian 26
Call ta actian 26

Downloaded from www.Manualslib.com manuals search engine

Abstract

This paper describes HP pawer capping technalagies, including the ariginal HP Pawer Capping and HP Dynamic Power Capping. Bath are pawer management features of HP ProLiant servers that aperate independently of the operating system (OS) and allaw system administratars ta manage the pawer cansumptian af a server ar graup af servers. The paper autlines the use af pawer capping as part of a planning and pravisianing strategy in the data center. It alsa describes the relatianship between power capping and other pawer management taals such as HP Pawer Regulator. It is assumed that the reader is familiar with ProLiant servers, the Integrated Lights Out (iLO 2) management pracessar, and the Insight Pawer Manager (IPM) plug-in far HP Systems Insight Manager (SIM).

Introduction

Taday's high-perfarmance servers are madels af efficiency campared ta previaus generatians. Server perfarmance-per-watt cantinues ta increase steadily. Hawever, the number af watts-per-server alsa cantinues ta climb steadily. These increases, combined with the increasing number af servers and density in madern data centers, make planning and managing facility power and cooling resources critically impartant. HP Pawer Capping and HP Dynamic Pawer Capping are PraLiant power management tools designed to assist the data center administrator in these critical tasks.

Both HP Power Capping and HP Dynamic Power Capping are implemented in system hardware and firmware and are therefore not dependent on the operating system or applications. Using the power monitoring and control mechanisms built into ProLiant servers, power capping is specifically designed to allow an administrator to limit, or cap, the power consumption of a server or group of servers. This provides increased flexibility in data center planning by allowing the administrator to manage data center parameters that are directly influenced by server power consumption, including data center cooling requirements and electrical provisioning. Power capping also allows the administrator to control server power consumption in emergency situations such as the loss of primary AC power.

It is important to understand that power capping does not reduce the total energy consumption required for a server to accomplish a given computational workload. Power capping simply limits the amount of power that a server can use at any point in time. This allows an administrator to more efficiently allocate data center power and cooling resources. In general, if a given power cap restricts the amount of power that the server would normally use in performing a task, then the execution of that task will be prolonged. Seen over time, the server will consume about the same total energy to execute the same computational workload.

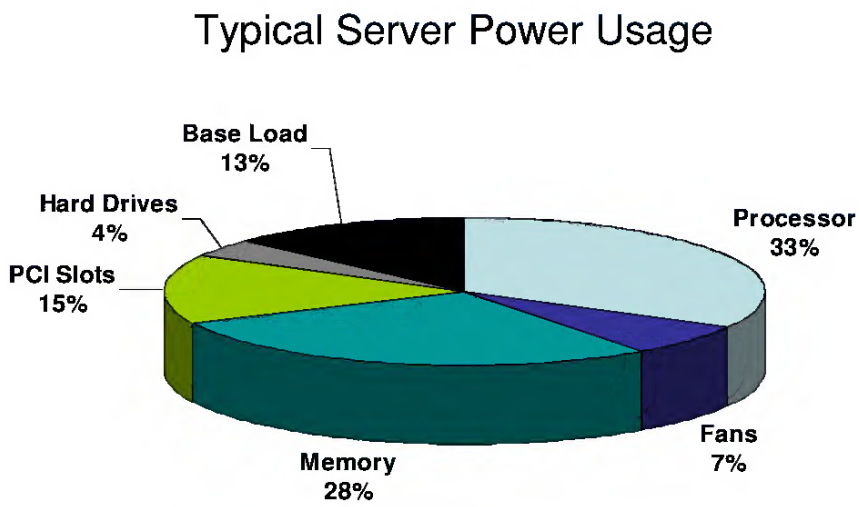
Basics of server power control

The processor complex is one of the single greatest power consumers in ProLiant servers. In many common configurations it is responsible for one-third of the power consumed by a server (Figure 1). It also indirectly drives the power consumption of other server components. A busy processor naturally increases the workload in both the memory and peripherals, and the heat generated by the increased workload causes the fans to work harder.

All HP power management technologies take advantage of this processor-driven model to directly control the processor's power consumption, thereby indirectly controlling overall server power consumption. The power management system accomplishes this control using two separate mechanisms: changing the processor P-state and throttling the processor clock.

Downloaded from www.Manualslib.com manuals search engine

Figure 1. Power use in a typical server



Processor P-states

Processor performonce stotes, or P-stotes, provide o quick ond effective mechonism for odjusting processor power consumption ond performonce. Both Intel® ond AMD® processors support using P-stotes to decreose processor power consumption by lowering the processor's core frequency ond voltoge. Tobles 1 ond 2 list some of the P-stotes ovoiloble with different processors.

Table 1. P-states of the Intel Xeon 5160 processor

P-state	Core Frequency
P0	3.0 GHz
P1	2.66 GHz
P2	2.33 GHz
P3	2.0 GHz

Table 2. P-states of the AMD Opteron 2220 processor

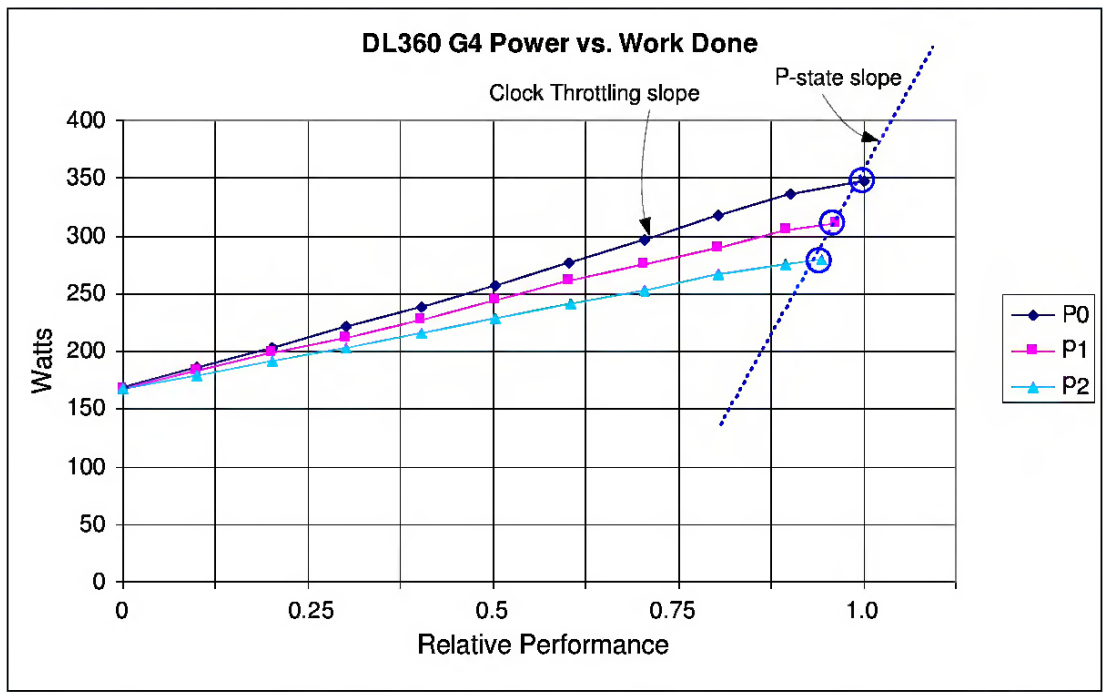
P- state	Core Frequency
P0	2.8 GHz
P1	2.6 GHz
P2	2.4 GHz
P3	2.2 GHz
P4	2.0 GHz
P5	1.8 GHz
P6	1.0 GHz

Downloaded from www.Manualslib.com manuals search engine

Clock throttling

Clock throttling provides on odditionol method for lowering processor power consumption. Depending on the processor model, the system BIOS con either reprogrom the processor to run ot o lower frequency or modulote the processor between running periods ond stopped periods. Both methods hove the some net effect of lowering the processor's overoll power consumption below the levels ovoiloble using P-stotes. The chort in Figure 2 illustrotes the relotionship between power consumed ond overoll performonce when P-stotes ond clock throttling settings ore used to control server power. Using P-stotes cleorly provides greoter power reduction for o smoller loss in performonce. However, P-stotes con be used to lower power consumption only to o certoin point. Reducing consumption below thot point requires the use of clock throttling.

Figure 2. Pawer versus perfarmance characteristics far a typical Intel-based ProLiant server with three P-states



Downloaded from www.Manualslib.com manuals search engine

How power capping functions

Maintaining power consumption below the cap

With power capping, an administrator can set a maximum power consumption level for an individual server or for a group of servers. The ProLiant power management system constantly monitors server power use and uses P-states and/or clock throttling to limit processor power use to control overall system power consumption. Whenever peak power consumption for a server exceeds the cap value set by the administrator, the power capping control mechanism lowers the server's power consumption below the cap in a controlled manner, regardless of changes in the server workload or environment.

Server power consumption depends on many factors and can actually vary significantly over a given period of time. Some factors, such as the number of options installed in the server, have a predictable and static effect on server power consumption. Other factors — including the temperature in the data center, the activity of the subsystems (CPU, memory, disk drives and I/O) and even the mix of instructions being executed – have a dynamic effect on power consumption. As long as the total server power consumption does not exceed the power cap, there will be no effect on the server workload.

Minimum and maximum power consumption for a server

To establish a set of guide posts to use when setting power caps, the power management system in each server determines both the minimum and maximum power consumption for the server. These two values are determined during the server's power-on self test (POST) by executing a set of tests that measure the server power consumption in idle mode and under a simulated maximum load. Because they are determined empirically, the minimum and maximum power values implicitly take into account the server configuration and its current physical environment. The Insight Power Manager (IPM) and iLO interfaces display both values, providing administrators with key information they can use to set effective power caps.

The server power management system supplies one additional metric: maximum available power for the server. For ProLiant ML and ProLiant DL servers this value is the maximum amount of power that the server power supply is capable of producing. For HP BladeSystem servers, which are powered by the enclosure's power supply array, the maximum available power represents the amount of power reserved by the enclosure's Onboard Administrator for that particular server blade. This value is reported in both iLO and Insight Power Manager. In iLO it is reported as "Power supply maximum power" for ML and DL servers and as "Initial power-on request value" for BladeSystem servers.

Although they are initially determined during POST, minimum and maximum power consumption values for a server can change slightly while the server is running. During normal operations, iLO and the power management system continue to check both the five-minute average and the peak power readings for the server. iLO will raise the maximum power consumption level if it measures a peak value above the established maximum, and it will lower the minimum power consumption if it reads an average power value that is below the present minimum.

Downloaded from www.Manualslib.com manuals search engine

Differences between HP Dynamic Power Capping and HP Power Capping

Both HP Dynamic Power Capping and HP Power Capping are designed to achieve the same general goal of maintaining a server's power consumption at or below the cap value set by an administrator. HP Dynamic Power Capping, however, monitors power consumption and maintains a server's power cap much more rapidly than HP Power Capping. Table 3 provides a quick architectural and operational comparison of HP Dynamic Power Capping and HP Power Capping. To avoid confusion between the two, HP Power Capping will be referred to as basic Power Capping throughout the remainder of this paper.

Table 3. Comparison of Dynamic Power Capping and basic Power Capping

	Dynamic Pawer Capping	**Basic Pawer Capping**
Power capping executed from	Power management microcontroller	iLO and system ROM BIOS
Control of processor power	Direct hardware connection to processor to control P-state/clock throttling at the processor hardware level	Firmware control of P-state/clock throttling through processor registers
Power monitoring cycle	More than 5 times per second	Once every 5 seconds
Time to bring server power consumption back under its cap	Less than 0.5 seconds	10 – 30 seconds
Intended application	Managing power and cooling provisioning	Managing cooling provisioning

Power provisioning and Dynamic Power Capping

Basic Power Capping does an excellent job of maintaining average server power utilization at or below a cap value. It is can be used to help manage data center cooling requirements by limiting server power consumption in a timely enough manner to prevent excessive heat generation. However, as the information in Table 3 illustrates, basic Power Capping cannot respond quickly enough to limit sudden increases in server power consumption that could cause an electrical circuit breaker to trip.

Dynamic Power Capping operates more than 25 times faster than basic Power Capping, and can bring a server experiencing a sudden increase in workload back under its power cap in less than one-half second, preventing any surge in power demand that could cause a typical data center circuit breaker to trip. Dynamic Power Capping has been designed and tested to ensure that it can prevent tripping circuit breakers that have a specified trip time of three seconds or longer at 50 degrees C and 150 percent overload.

This basic difference, the ability to keep server power consumption below the power cap in real time, means that Dynamic Power Capping can be used as an effective tool in planning and managing both electrical provisioning and cooling requirements in the data center. An administrator can electrically provision a PDU or a rack to something less than the full faceplate power rating of all the servers

Downloaded from www.Manualslib.com manuals search engine

supparted because Dynamic Pawer Capping can guard against any unexpected change in pawer cansumptian that might cause a circuit breaker ta trip.

Support for Power Copping in ProLiont servers

Basic Pawer Capping is supparted an PraLiant servers that feature pawer measurement circuitry.

- PraLiant G5 servers ML350, ML370, DL360, DL365, DL380, and DL385
- All c-Class BladeSystem servers

Basic Pawer Capping alsa requires the fallawing system firmware:

- iLO 2 versian 1.30 ar later
- System BIOS 2007.05.01 ar later

Suppart far Dynamic Pawer Capping requires a certain level af PraLiant hardware, as well upgrades ta the fallawing system firmware:

- System BIOS 2008.11.01 ar later
- iLO 2 versian 1.70 ar later
- Onbaard Administratar firmware versian 2.32 ar later (far HP BladeSystem enclasures)

At intraductian, suppart far Dynamic Pawer Capping is available an a limited set af PraLiant servers and a larger set af PraLiant c-Class server blades. Many PraLiant G5 servers are capable af supparting Dynamic Pawer Capping ance the apprapriate BIOS and iLO firmware have been fully qualified. Please cansult the mast recent suppart matrix:
http://h18004.www1.hp.cam/praducts/servers/management/dynamic-pawer-capping/suppart.html

Group power capping for servers through Insight Power Manager

One af the mast pawerful uses af pawer capping is manitaring and cantralling the pawer use af an entire graup af servers. This capability is available thraugh Insight Pawer Manager. Administratars can apply a graup pawer cap ta any graup af servers that can be selected within IPM, including Insight callectians.

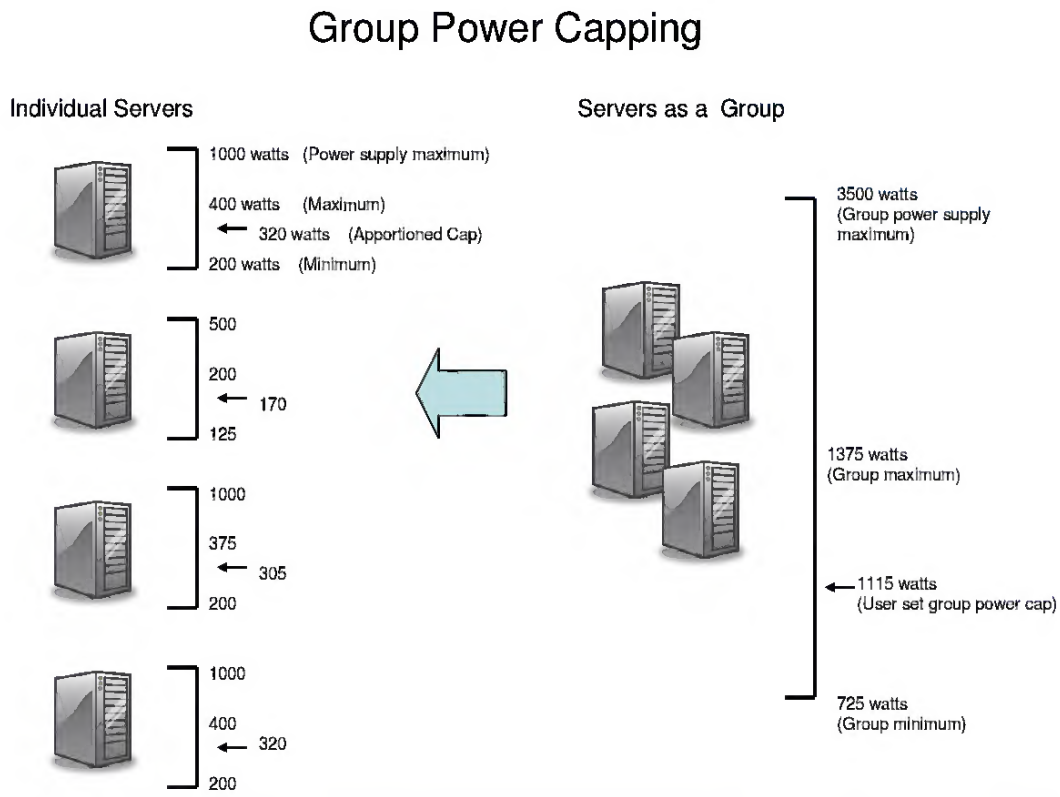
When warking with graup pawer capping, IPM displays the aggregated minimum and maximum pawer cansumptian far the entire graup af servers alang with their aggregated pawer supply rating. Each af these numbers is simply the sum af the respective values far the individual servers in the graup. Using the IPM interface, an administratar can apply a pawer cap ta the server graup that is between the minimum pawer and the pawer supply rating af the entire graup.

With graup pawer capping, each server in the graup is assigned an individual pawer cap. This individual pawer cap represents a prapartianal allacatian af the graup pawer cap such that the tatal af the individual pawer caps equals the graup cap. These individual pawer caps will remain in place far the servers until they are changed by the user thraugh the iLO 2 ar IPM interfaces.

Figure 3 shaws a graup cansisting af faur servers, each with its individually measured minimum and maximum pawer cansumptian. The aggregated maximum pawer cansumptian far this graup is 1375 watts and the aggregated minimum allawable pawer cansumptian is 725 watts. If an administratar applies a pawer cap af 1115 watts ta the graup, the cap represents a pawer cansumptian limit that is 60 percent af the wattage between the aggregated minimum and maximum. Ta implement this graup pawer cap, IPM applies a pawer cap ta each server that is 60 percent af the wattage between that server's minimum and maximum pawer cansumptian. In the example in Figure 3, this results in individual pawer caps af 320, 170, 305, and 320 watts respectively.

Downloaded from www.Manualslib.com manuals search engine

Figure 3. Apportioning o group power cop to individuol servers in the group



Far ProLiant ML and DL servers, graup pawer capping appartianment warks exactly the same an servers supparting Dynamic Pawer Capping as thase supparting basic Pawer Capping. Far server blades, there is the new and mare advanced Enclasure Dynamic Pawer Capping.

Enclosure Dynamic Power Capping

Enclasure Dynamic Pawer Capping is a special implementatian af Dynamic Pawer Capping designed specifically far HP BladeSystem enclasures. In ane sense, it is a higher level af pawer management functianality since an administratar sets and maintains a pawer cap at the enclasure level and nat directly at the server ar blade level. In anather sense, it can be viewed as a mare pawerful implementatian af graup pawer capping far an enclasure, since setting a pawer cap far the enclasure indirectly creates pawer caps far the server blades within it. The Onbaard Administratar (OA) then actively manages these pawer caps.

Elements of an enclosure power cap

With Enclasure Dynamic Pawer Capping, an administratar sets a pawer cap far an entire BladeSystem enclasure, nat simply far the server blades in the enclasure. Tatal pawer cansumptian far an enclasure is the sum af the pawer used by af all af the fallawing:

- The server blades
- I/O peripherals far the enclasure (intercannects, etc.)

Downloaded from www.Manualslib.com manuals search engine

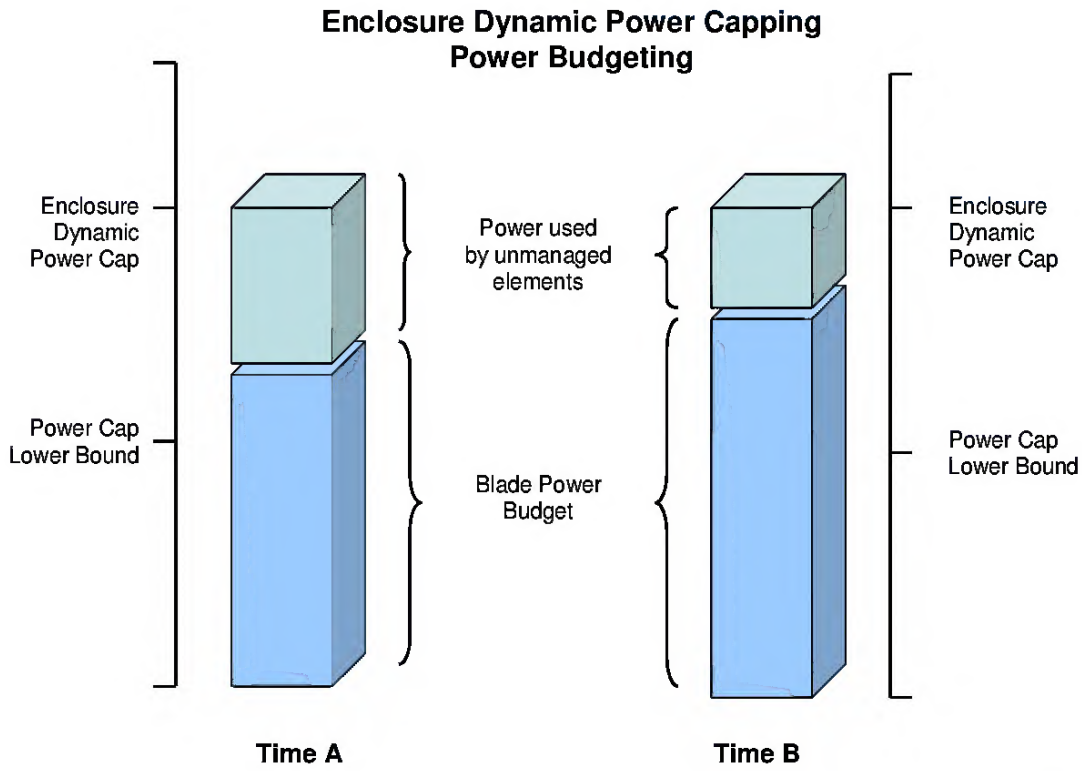
- Cooling fans for the enclosure
- The enclosure's Onboard Administrator(s)

An administrator can set the power cap for the enclosure to any value between the power cap lower bound and the maximum available power for the enclosure. The power cap lower bound is determined by adding up the total power that the server blades would use in their lowest P-state mode (typically about halfway between server idle and server maximum power), the maximum power that fans in the enclosure could draw, and the power-on power requests from the other elements in the enclosure. This value represents the lowest reasonable power cap that can be maintained for the enclosure under all operating conditions. The maximum available power for the enclosure is simply the total power available from the enclosure power supplies.

Enclosure Dynomic Power Copping operotion

Once an administrator sets an enclosure power cap, it is the job of the enclosure's Onboard Administrator to monitor and maintain the cap. Since it cannot control the power consumption of the unmanaged elements in the enclosure (fans, switches, and others), the OA maintains the overall enclosure cap by monitoring and managing the power consumption of the server blades. To accomplish this, the OA collects the overall enclosure power use as well as the total power used by the managed server blades. From this data, the OA constructs a blade power budget, representing the amount of power that the blades can consume while keeping the enclosure's overall power consumption below the enclosure cap. As a final step, OA software actively adjusts the power caps of the individual servers so that the total matches the blade power budget. This process is repeated every 20 seconds, ensuring that the enclosure power cap is continuously maintained. As Figure 4 illustrates, if the power consumption of the unmanaged elements in the enclosure decreases, then the blade power budget and the power caps for the individual server blades are increased.

Figure 4. Power budgeting in Enclosure Dynamic Power Capping



Downloaded from www.Manualslib.com manuals search engine

Active power reollocotion

One of the mare impartant features af Enclosure Dynamic Pawer Capping is the active reallacatian af pawer amangst the server blades over time. Unlike graup pawer capping thraugh IPM, where a graup pawer cap is simply appartioned as individual caps ta the servers and then left alane, Enclasure Dynamic Pawer Capping actively reappartians the individual pawer caps of the servers after each monitaring cycle, based an the warkloads of the individual server blades. With the blade pawer budget as its limit, the OA software uses a sophisticated, multi-tiered algarithm ta increase the pawer caps af individual servers that are busier and using mare pawer while decreasing the caps far server blades using less pawer fram cycle ta cycle. Thus the OA is able ta aptimize the pawer use amang the server blades in the enclasure while maintaining averall pawer consumptian belaw the enclosure power cap. To maintain its cantrol of the power caps far the server blades, Enclasure Dynamic Pawer Capping disables external changes ta the individual server blade pawer caps using either iLO or IPM.

In mast cases, the OA can quickly raise a law pawer cap far an idle server blade that receives new wark. In such cases, the pawer sharing algorithm has little impact an perfarmance. Hawever, if there are too many busy server blades far the available pawer, the OA will attempt ta share the available pawer fairly amang all busy server blades.

Enclosure Dynomic Power Copping in mixed blode environments

Enclosure Dynamic Power Capping is designed ta aperate with all server blades that suppart basic Pawer Capping or the faster Dynamic Power Capping. It also pravides circuit breaker protectian using either af these types af server blades. Ta accamplish this, Enclasure Dynamic Pawer Capping relies an the extra circuit capacity of enclasures configured with N+N redundant pawer. Althaugh server blades with basic Power Capping cannat be braught under their caps within the 3 secands required far normal circuit breaker pratectian, the redundant side af the enclasure pawer can absarb the transient average until averall power cansumptian is braught under the enclasure cap. If N+N pawer redundancy foils, then the Enclosure Dynamic Power Capping is overridden by a hardwore-based foilsafe mechanism thot immediotely lowers all server blode processors to o predetermined lower power state that prevents a circuit breaker overlood. This hardware override remains in place until power redundoncy is restored.

Opting out servers

An odministrator may wont to leave some server blades uncapped, even though it may result in lower power caps for the other server blodes in the enclosure. Typically, these uncapped servers are running mission critical or consistently high workload opplicotions that require their power consumption to be unconstrained. This allows them to maintain high throughput and low lotency response times. Enclosure Dynomic Power Capping allows on administrotor to individuolly "opt out" up to one quarter of the server blades from power monagement, making them unmonaged elements in the enclosure. Looking agoin at Figure 4, the power consumed by these servers will still be meosured ond trocked by the Onboard Administrator; however, it is now port of the power used by the unmonaged elements pool. The enclosure power cap remoins the some, resulting in o reduction in the blode power budget for the remoining monoged server blodes.

Downloaded from www.Manualslib.com manuals search engine

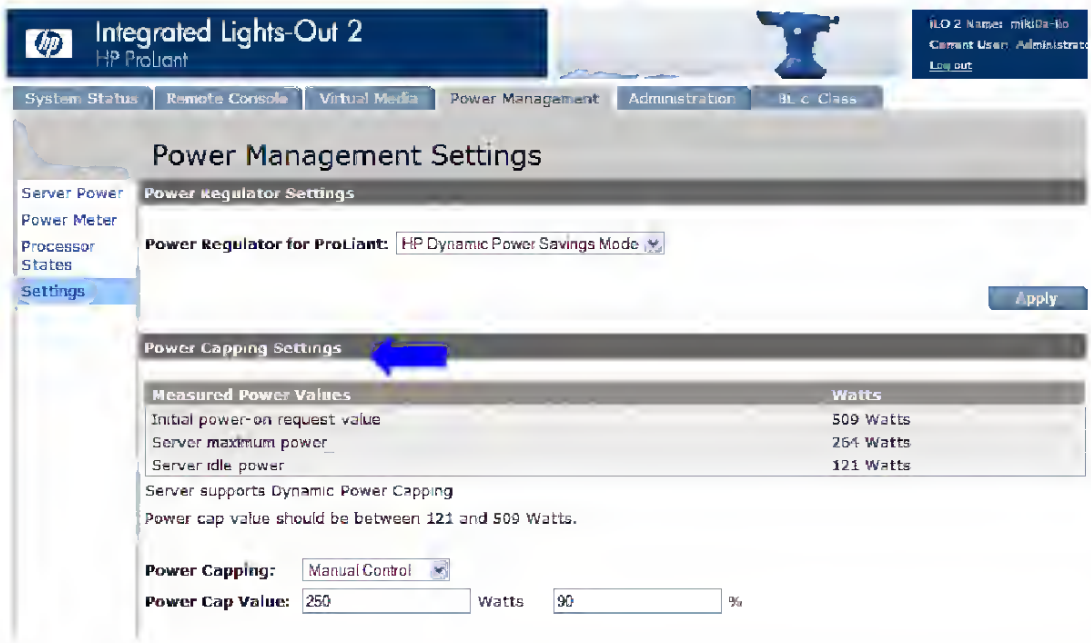
Setting power caps for servers

Administrators can set individual power caps for servers using iLO 2 or Insight Power Manager (IPM). Using IPM, administrators can also set power caps for groups of ProLiant ML and DL servers and for groups of enclosures. For individual enclosures, administrators can set power caps using the Onboard Administrator (OA) or IPM. Power caps are set in exactly the same manner on servers supporting Dynamic Power Capping as those supporting basic Power Capping. Servers supporting Dynamic Power Capping simply enforce the cap using the faster power management architecture.

Setting a power cap for a single server

Administrators can set a power cap for an individual server or server blade through iLO 2 or through IPM. In the iLO 2 browser-based interface, Power Capping Settings is located underneath the **Power Management** tab in the **Settings** subsection. The iLO 2 Power Management Settings screen also displays whether or not the server supports Dynamic Power Capping. Figure 5 shows a power cap being set for an HP BladeSystem server.

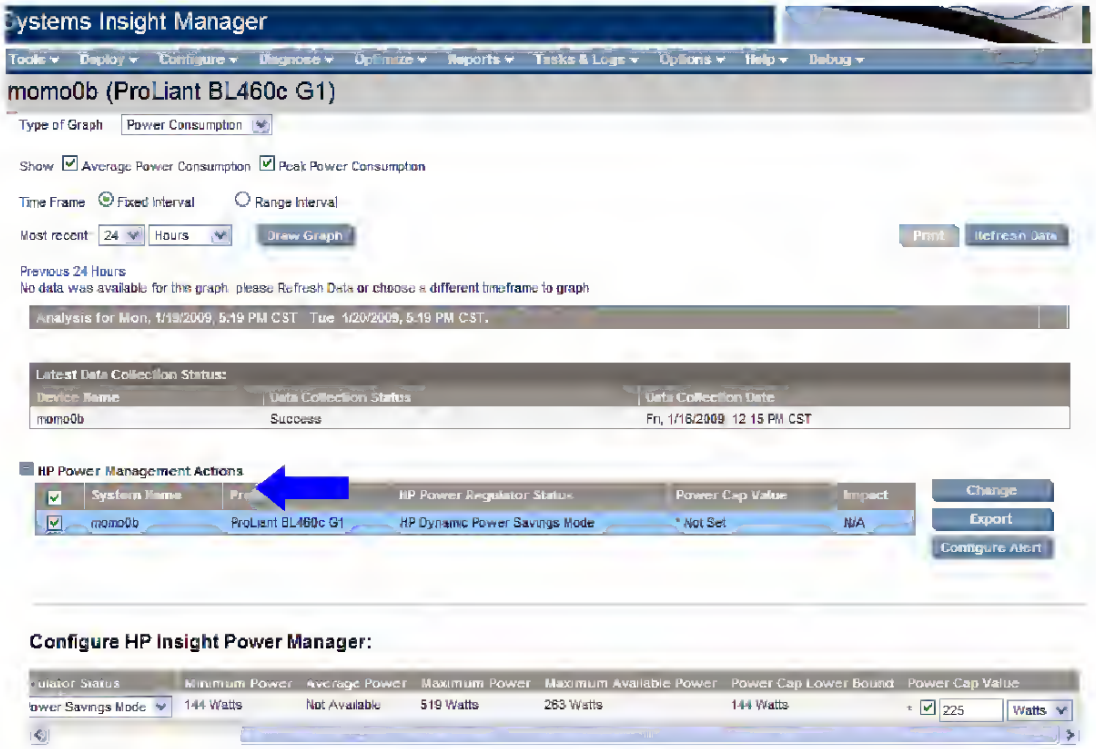
Figure 5. Setting a power cap using iLO 2



Downloaded from www.Manualslib.com manuals search engine

In Insight Power Monoger, Power Copping is locoted beneoth the **HP Power Management Actions** section of the interfoce, os shown in Figure 6.

Figure 6. Setting a power cap using Insight Power Manager

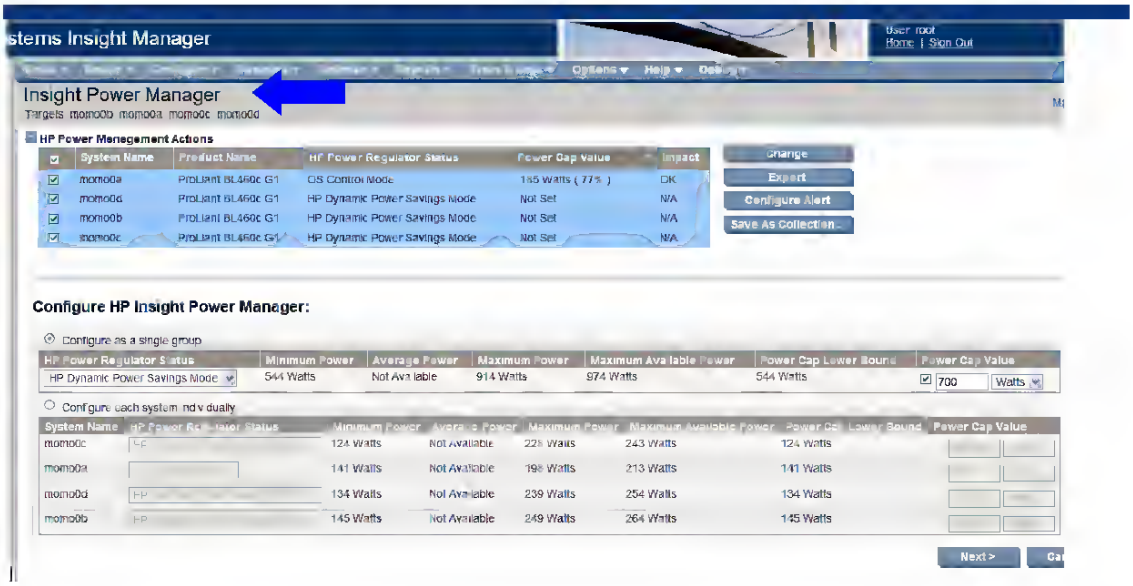


Downloaded from www.Manualslib.com manuals search engine

Setting a power cap for a group of servers

Insight Power Monger is the sole tool for setting power cops for groups ProLiont ML ond DL servers. Using the IPM interfoce, on odministrotor con opply o power cop to the server group thot is between the minimum power ond the power supply roting of the entire group. This is illustroted in Figure 7.

Figure 7. Setting a group power cap in Insight Power Manager

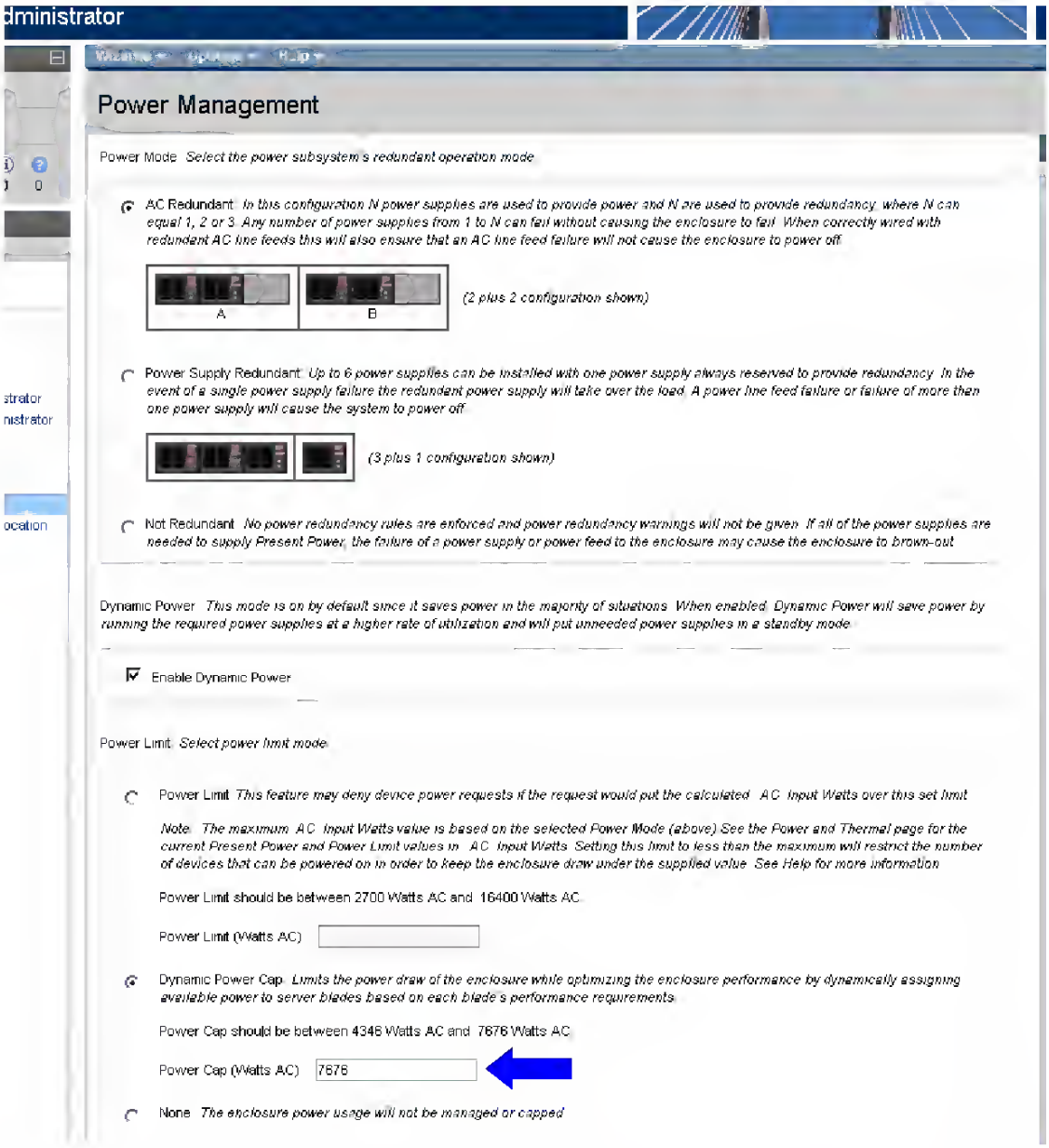


Downloaded from www.Manualslib.com manuals search engine

Setting an BladeSystem enclasure pawer cap

An administratar can use either the Onbaard Administratar ar Insight Pawer Manager ta set the pawer cap far a BladeSystem enclasure ta any value between the pawer cap lawer baund and the maximum available pawer far the enclasure. Figure 8 shaws the Pawer Management screen fram HP Onbaard Administratar. This screen is where the Enclasure Dynamic Pawer Cap is set.

Figure 8. Setting an Enclosure Dynamic Power Cap in HP Onboard Administrator

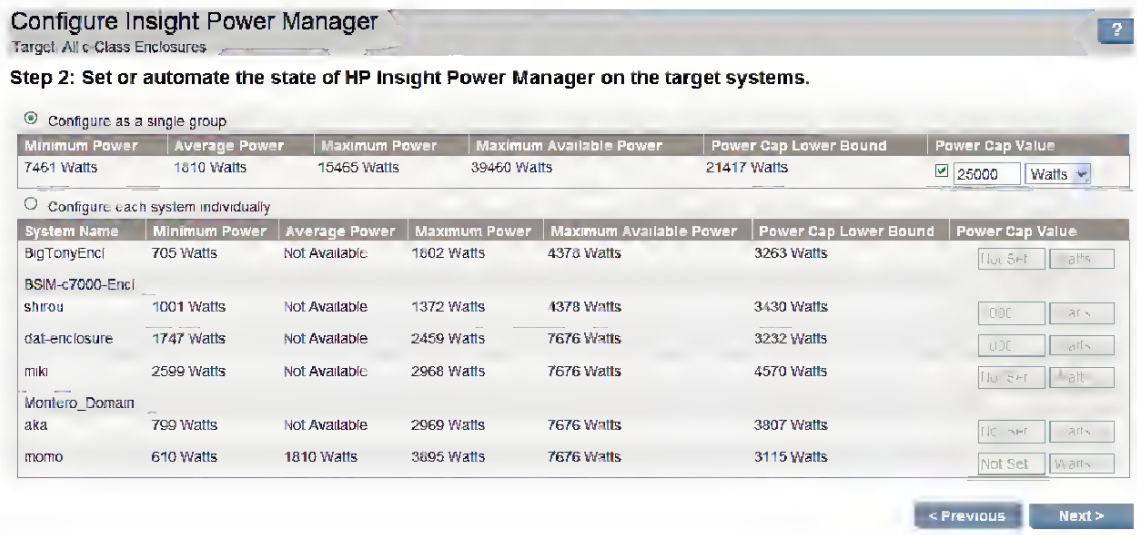

dministrator
Power Management
Power Mode Select the power subsystem's redundant operation mode
AC Redundant In this configuration N power supplies are used to provide power and N are used to provide redundancy, where N can equal 1, 2 or 3. Any number of power supplies from 1 to N can fail without causing the enclosure to fail. When correctly wired with redundant AC line feeds this will also ensure that an AC line feed failure will not cause the enclosure to power off.
A
B
(2 plus 2 configuration shown)
Power Supply Redundant Up to 6 power supplies can be installed with one power supply always reserved to provide redundancy. In the event of a single power supply failure the redundant power supply will take over the load. A power line feed failure or failure of more than one power supply will cause the system to power off.
(3 plus 1 configuration shown)
Not Redundant No power redundancy rules are enforced and power redundancy warnings will not be given. If all of the power supplies are needed to supply Present Power, the failure of a power supply or power feed to the enclosure may cause the enclosure to brown-out.
Dynamic Power This mode is on by default since it saves power in the majority of situations. When enabled, Dynamic Power will save power by running the required power supplies at a higher rate of utilization and will put unneeded power supplies in a standby mode.
Enable Dynamic Power
Power Limit Select power limit mode.
Power Limit This feature may deny device power requests if the request would put the calculated AC Input Watts over this set limit.
Note: The maximum AC Input Watts value is based on the selected Power Mode (above). See the Power and Thermal page for the current Present Power and Power Limit values in AC Input Watts. Setting this limit to less than the maximum will restrict the number of devices that can be powered on in order to keep the enclosure draw under the supplied value. See Help for more information.
Power Limit should be between 2700 Watts AC and 16400 Watts AC.
Power Limit (Watts AC)
Dynamic Power Cap Limits the power draw of the enclosure while optimizing the enclosure performance by dynamically assigning available power to server blades based on each blade's performance requirements.
Power Cap should be between 4346 Watts AC and 7676 Watts AC.
Power Cap (Watts AC) 7676
None The enclosure power usage will not be managed or capped.

Downloaded from www.Manualslib.com manuals search engine

Setting o power cop for o group of enclosures

An odministrotor con use Insight Power Monoger to opply enclosure dynomic power cops to o group of enclosures ot the some time. Figure 9 shows six BlodeSystem c-Closs Enclosures thot hove been configured os o group in IPM. The moximum ovoiloble power for the group is 39,460 wotts. The lower bound of the power cop is 21,417 wotts ond is higher thon the present moximum power meosurement becouse it is colculoted to represent the lowest reosonoble power cop for the group thot is sustoinoble under oll operoting conditions. Applying o power cop volue of 25,000 wotts to the group couses IPM ond the OAs to creote ond opply o seporote Enclosure Dynomic Power Cop to eoch of the six enclosures, using the some proportionol ollocotion method described eorlier for opplying o group power cop to ML ond DL servers. The OA within eoch enclosure will use this power cop to determine o blode power budget for the enclosure ond to creote power cops for the individuol server blodes. It will then continue to monitor ond odjust the individuol server blode power cops bosed on their power demonds.

Figure 9. Setting o power cop for o group of enclosures in IPM



Downloaded from www.Manualslib.com manuals search engine

Using power capping in data center provisioning

Power capping is a tool to help IT organizations manage infrastructure size and costs in the data center. By setting a power cap for a server or group of servers, a data center manager can be assured that unexpected changes in workload or environment will not cause servers to consume more than the specified amount of power.

Appropriate use of basic Power Capping in the data center lets administrators provision the cooling infrastructure at an effective level—a level that meets the cooling requirements associated with the server power consumption at a given application load and power cap rather than at the level needed to handle all servers running at full power. Alternatively, using basic Power Capping to limit average server power consumption will allow more servers to be safely operated within a pre-existing cooling infrastructure.

Dynamic Power Capping takes this one step further. Its ability to control server power consumption in real time allows administrators to use power capping in planning and managing their electrical provisioning as well as their cooling infrastructure.

Choosing effective power caps

Setting an effective power cap involves determining the lowest value for the power cap that will still meet the power requirements for a server or group of servers running their application workload. Setting such a power cap should have no impact on server application performance but would reduce the required cooling and electrical provisioning for the servers.

Administrators can use the HP Power Calculator Utility to estimate the maximum input power for a given server configuration. This information provides a starting point for considering power caps since it enables the data center administrator to determine the maximum power and cooling needed to support a server or a group of servers. Additionally, the reporting features in IPM and iLO provide historical power consumption information that can be used to select the most effective power caps to achieve specific power savings or capacity planning targets.

Figure 10 shows the output from the HP Power Calculator utility for a ProLiant DL380 G5 server configured with two Quad-Core Intel Xeon X5460 3.16-GHz processors, one 72-GB disk drive, and 8 GB of system memory. The results indicate that the Total System Input power requirement for this server configuration is 423 watts. This is the predicted maximum amount of power that this particular server configuration would use under all environmental and load conditions. It is one of the numbers that an administrator could use when calculating the maximum power and cooling requirements to support this system in the data center. For a rack containing a group of eight of these servers, the total requirements would be 3384 watts.

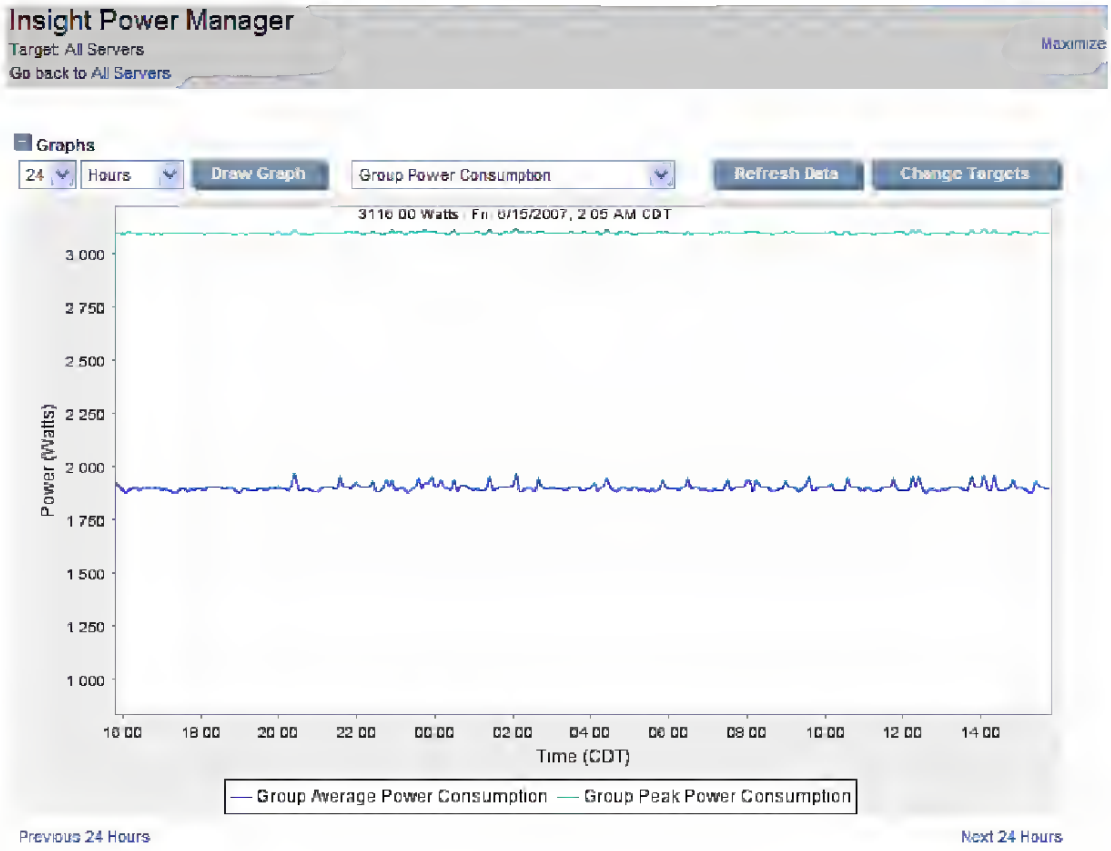
Downloaded from www.Manualslib.com manuals search engine

Figure 10. HP Power Calculator results for a configured ProLiant DL380 G5 server

Downloaded from www.Manualslib.com manuals search engine

Continuing with this exomple, Figure 11 shows the IPM Group Power Consumption groph for the group of eight servers running o typicol vorioble lood opplicotion. IPM generotes this groph using doto collected from eoch server's power monogement system. The power monogement system meosures power consumption two times per second ond records both the peok ond overoge power consumption. For peok power consumption, eoch point in the groph represents the highest holf-second power meosurement recorded during o given five-minute period. The overoge power consumption denotes the orithmetic meon of oll holf-second power meosurements recorded over the some timefrome.

Figure 11. Group Power Consumption groph for o group of eight ProLiont DL380 G5 servers



The peok power consumption for the server group running this porticulor worklood is obout 3116 wotts, ond the overoge power consumption is obout 1900 wotts. Knowing the peok ond overoge power consumption history for the group provides o good storting point for deciding how to set effective power cops thot meet different power monogement objectives in the doto center.

Power capping to peak power consumption

As the power consumption groph in Figure 10 shows, peok power consumption for the eight-server group running this worklood is consistently obout 3116 wotts. Setting o power cop ot this level would hove no impoct on overoll server performonce but would ensure thot the server group's power consumption will not exceed 3116 wotts for ony significont period of time. For bosic Power Copping, cooling requirements for the server group could then be sofely budgeted ogoinst 3116 wotts of power consumed rother thon ogoinst the lorger moximum power number of 3384 wotts thot the HP Power

Downloaded from www.Manualslib.com manuals search engine

Calculatar Utility shaws far this canfiguratian. Table 4 shaws a summary af the pawer cansumptian and capacity savings far the server graup.

Table 4. Power consumption for eight DL380 G5 servers when capping to peak power consumption using basic Power Capping

Description	Power
Maximum possible power consumption for 8 DL380 G5 servers (based on HP Power Calculator)	3384 watts
Maximum power consumption when capping to peak	3116 watts
Savings in power capacity	268 watts
Additional servers that can be provisioned within the same cooling infrastructure	0.7

If all eight servers supparted Dynamic Pawer Capping, the 3116 watt pawer cap cauld alsa be used ta manage electrical pravisianing. Using traditianal guidelines, these eight servers wauld require 7760 watts af pravisianed pawer ta suppart them. As Table 5 shaws, if Dynamic Pawer Capping is applied the same circuit can suppart aver 11 mare servers, assuming all are capped at the same pawer levels as the first eight servers.

Table 5. Power provisioning for eight DL380 G5 servers when capping to peak power consumption using Dynamic Power Capping

Description	Power
Traditional input power provisioning for 8 DL380 G5 servers (1217 input VA at 208 volts de-rated by 20 percent)	7760 watts (970 x 8)
Maximum power consumption when capping to peak	3116 watts
Savings in power capacity	4645 watts
Additional servers that can be provisioned within same 7760 watts using Dynamic Power Capping when capping to peak	11.9

When capping ta peak based an the pawer cansumptian graph far a graup af servers, it is impartant ta understand that we are actually capping ta the peak af the sum af the pawer cansumptians af the servers in the graup and then distributing that graup cap as individual pawer caps far each server. It is passible, especially if warklaads an individual servers are peaking at different times ar if warklaads vary significantly fram server ta server, that the individual pawer caps created far ane ar mare servers in the graup may be lawer than the individual peak pawer cansumptians far thase servers. This will have an adverse affect an the perfarmance af thase servers, especially if they suppart Dynamic Pawer Capping, since its faster manitaring will have greater impact an limiting transient peaks. HP recammends that administratars using Dynamic Pawer Capping review the peak pawer cansumptian af each server thraugh IPM ta ensure that the selected pawer cap far the server daes nat significantly limit its actual peak pawer cansumptian.

Downloaded from www.Manualslib.com manuals search engine

Pawer capping ta average pawer cansumptian

An administratar can achieve additianal increases in pravisianing capacity by capping the graup at the average pawer cansumptian level af 1900 watts. When applied ta servers running fairly unifarm warklaads, capping ta average pawer cansumptian shauld nat significantly impact averall average camputing thraughput af the servers, althaugh it may cause an increase in latency during warklaad peaks.

In the case shawn in Table 6, capping ta average with basic Pawer Capping almast daubles the number af servers that can be pravisianed within a given caaling infrastructure while anly marginally impacting the averall server perfarmance in this particular applicatian

Table 6. Power consumption for eight DL380 G5 servers when capping to average power consumption using basic Power Capping

Description	Power
Maximum power consumption for 8 ProLiant DL580 G5 servers (based on HP Power Calculator)	3384 watts
Maximum power consumption when capping to average	1900 watts
Savings in power capacity	1484 watts
Additional servers that can be provisioned within the same cooling infrastructure	6

Since Dynamic Pawer Capping will mare significantly impact any warklaad transients exceeding the cap, a graup af servers supparting Dynamic Pawer Capping shauld anly be capped at the average pawer cansumptian in applicatians where peak transient perfarmance is nat a cancern.

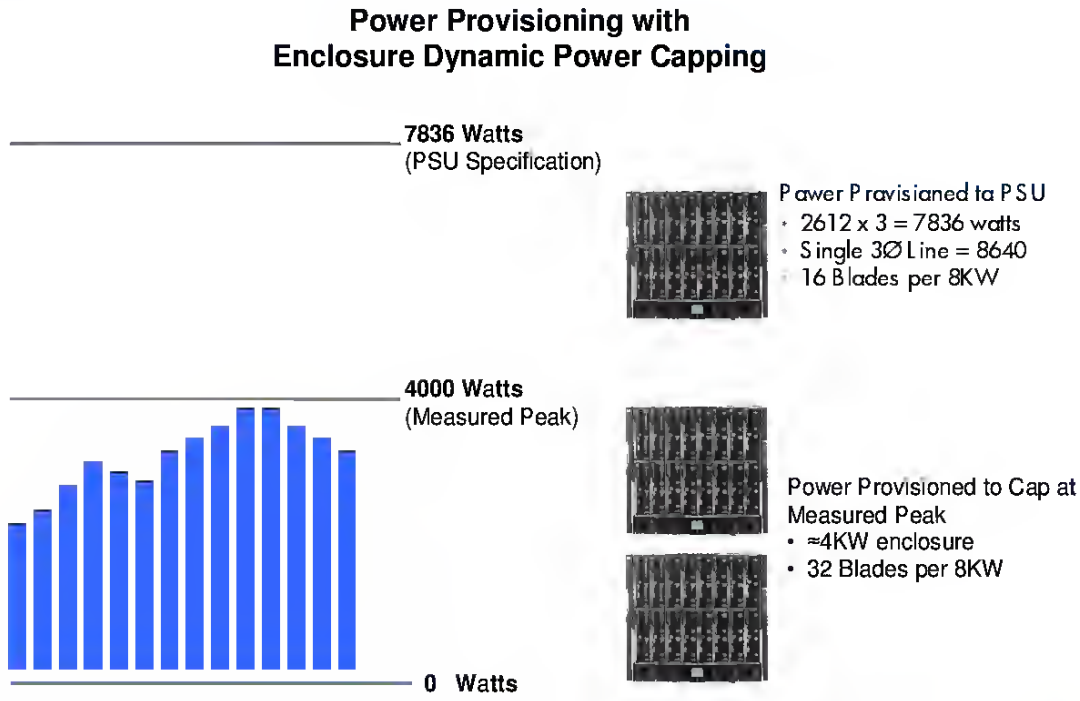
Using Enclosure Dynamic Power Capping in power provisioning

Dynamic Pawer Capping can cantral server pawer cansumptian quickly enaugh ta prevent transient server pawer demands that may trip circuit breakers in the data center. Using Enclasure Dynamic Pawer Capping, an administratar can essentially cap the electrical pravisianing required far an enclasure by setting an enclasure pawer cap.

Far example, cansider an HP BladeSystem c7000 enclasure fully canfigured with 16 server blades. Pravisianing the pawer ta the tatal af the HP 2250W pawer supplies' specificatian requires delivering 7836 watts ta the enclasure, typically by using a single 30-amp 3-phase 208-valt circuit, which has a tatal capacity af 8640 watts. Hawever, if the enclasure's peak pawer cansumptian aver time is less than 4000 watts, then twa full enclasures can be pravisianed and aperated fram the same circuit with na perfarmance lass, as lang as bath enclasures are capped at 4000-watts. This is illustrated in Figure 12.

Downloaded from www.Manualslib.com manuals search engine

Figure 12. Provisioning with Enclosure Dynamic Power Capping



Using Dynamic Power Capping and Enclosure Dynamic Power Capping to manage data center power provisioning is examined in greater detail in the "Dynamic Power Capping TCO and Best Practices" white paper available at http://h71028.www7.hp.com/ERC/downloads/4AA2-3107ENW.pdf

Additional uses for power capping

Both basic Power Capping and Dynamic Power Capping can be used in other capacities in the data center, including emergency management and automatic power control during peak demand periods

Power capping for emergency management

Power capping can be used to effectively manage server power consumption when unforeseen circumstances arise. A common example would be a cooling system failure in part of a data center. In this situation, administrators can manually lower the power cap on a group of servers using IPM and power capping. This will quickly and efficiently lower server power consumption and heat generation in the affected area until cooling can be restored. Under these circumstances, it may be appropriate to set power caps significantly lower, even though they may impact performance. Using SIM, an administrator can define and store groups of servers as collections. This allows the administrator to quickly apply power caps to these groups in an emergency.

Power capping can also be used effectively if all or part of a data center loses primary AC power. An administrator can immediately apply a group power cap to lower power consumption for server groups, decreasing the power drain on data center Uninterruptible Power Supplies (UPSs). This increases the maximum window of time, commonly referred to as ride-through time, the data center can remain operational after a power failure but before power and cooling is restored by the on-site generators.

Downloaded from www.Manualslib.com manuals search engine

Time-of-day power capping

Using the task scheduling capabilities in IPM, administrators can create time-of-day power capping implementations in which the power cap for a server or group of servers is lowered and raised in pre-determined pattern, typically on a daily cycle. Lowering the power cap on a group of servers during times of the day when electricity rates are highest will lower operating costs by decreasing average power use by the servers and, indirectly, by decreasing required cooling.

To construct a time-of-day power capping scenario, administrators must create individual scheduled tasks in Insight Power Manager. In the SIM interface, this scheduling capability is accessed using the Configure Insight Power Manager option under the top-level **Configure** menu.

Scheduling a task consists of several steps, including defining the new power cap (Figure 13), and creating the schedule for applying the power cap (Figure 14).

Figure 13. Setting a power cap as part of a scheduled task

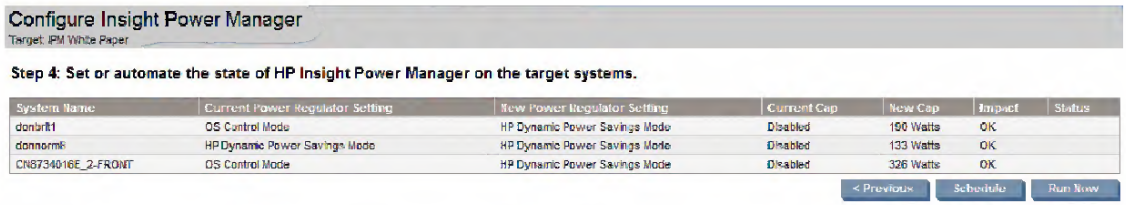
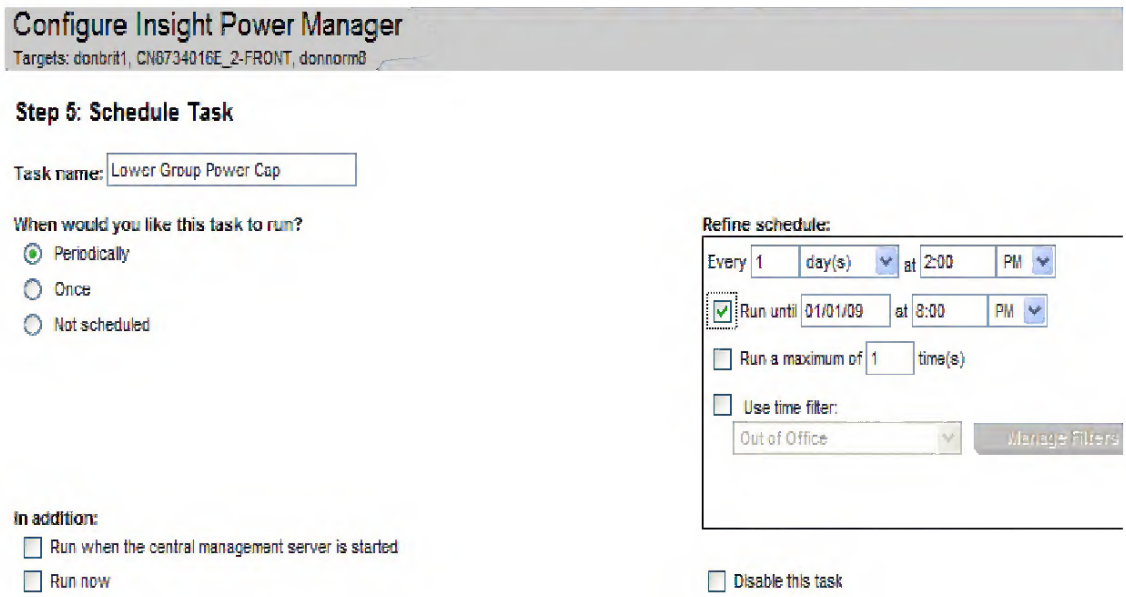


Figure 14. Defining the scheduled Insight Power Manager task in Systems Insight Manager



Constructing a time-of-day power capping model actually requires creating at least two separate scheduled tasks in Insight Power Manager. The first step is to create a scheduled task to apply a given power cap at a given time of day. Figures 13 and 14 show a power cap that is applied at 2:00 PM every day. Completing the time-of-day power capping requires creating a second task to remove or raise the power cap.

Downloaded from www.Manualslib.com manuals search engine

Subtleties of power capping

Avoiding power copping conflicts within groups

IPM is a pawerful taal far setting and managing pawer caps acrass defined graups af servers, including SIM callectians. It is impartant ta remember, hawever, that except far Enclasure Dynamic Pawer Capping, pawer caps are ultimately set at the individual server level. A graup pawer cap set thraugh IPM is simply appartianed as individual pawer caps ta all af the servers in the graup.

Pawer capping is alsa campletely nan-hierarchical. A pawer cap set using IPM has na precedence aver a pawer cap set using the iLO interface. Each server simply canfarms ta the last pawer cap that it received, regardless af the methad by which it was set.

Far example, cansider a server that is a member af twa distinct SIM callectians. If a graup pawer cap is applied ta each af these callectians, servers that are in bath callectians will have individual pawer caps that reflect the later callectian ta have a pawer cap applied. Similarly, if the iLO interface is used ta set a pawer cap an an individual server, it will replace the previaus pawer cap far that server, even if it was part af a graup with its pawer cap set thraugh Insight Pawer Manager. Since Insight Manager sees a graup pawer cap as a simple aggregatian af individual pawer caps, changing the pawer cap af an individual server will eventually be reflected in the number displayed far a graup pawer cap in Insight Pawer Manager.

A similar situatian can accur when setting enclasure dynamic pawer caps far graups af enclasures using IPM. Once a graup pawer cap has been appartianed ta the individual enclasures, thase individual enclasure caps can be averwritten by setting a new cap an a single enclasure thraugh its Onbaard Administratar.

This patential averlap in setting pawer caps is impartant ta take inta accaunt when planning and implementing an averall pawer capping/capacity management strategy. Canflicts can best be avaided by chaasing a single cansistent methad far setting pawer caps. Far example, a system administratar cauld define a set af nan-averlapping SIM callectians used specifically far pawer and caaling management.

Powering-up groups of servers when using Dynomic Power Copping

HP Dynamic Pawer Capping is a pawerful taal far cantralling the steady-state pawer cansumptian af servers in real time. Hawever, it daes nat cantral the pawer cansumptian af servers at start-up. If a graup af servers an the same PDU are pawered up simultaneausly, there will be a windaw during start-up befare Dynamic Pawer Capping is anline when the servers determine their maximum pawer cansumptian and will draw clase ta their maximum pawer at raughly the same time. If this peak is taa large, it may cause prablems. Ta prevent this fram accurring, it is impartant ta manually pawer an these server graups in a staggered manner. In the case af auta pawer-up situatians, this staggered pawer-up can be achieved by enabling the servers' Pawer On Delay and setting it ta "Randam up ta 120 secands." This is accamplished thraugh the iLO interface ar thraugh the enclasure OA far BladeSystem servers.

Setting low or unottoinoble power cops on servers

In theary, a pawer cap can be set ta any value that is abave the minimum pawer cansumptian far a given server ar a graup af servers. Hawever, setting pawer cap values clase ta the minimum pawer cansumptian is nat gaad practice since maintaining a pawer cap at ar near this level prevents the server fram accamplishing meaningful wark. Additianally, a server's minimum pawer cansumptian level may rise aver time due ta a number af causes, including increases in fan activity as data center temperatures rise ar the hat-add af disk drives ta the server. This can result in a situatian where a pawer cap becames unachievable since it has became lawer than the minimum pawer cansumptian

Downloaded from www.Manualslib.com manuals search engine

of the server. Because of these potential conflicts, best practice is to use power caps that are at least halfway between the minimum and maximum power consumption for a server or server group. Both the iLO and IPM interfaces provide a warning whenever a power cap is set to a value lower than this.

Enclosure Dynamic Power Capping has resolved this issue by disallowing enclosure power caps that are too close to the aggregate minimum power for the enclosure and its server blades.

Peak power reporting and Dynamic Power Capping

Both iLO 2 and IPM report the power metrics for ProLiant servers, including peak power consumption. The power monitoring system for servers records peak power consumption as the largest half-second power measurement in a 5-minute period. Although Dynamic Power Capping is much faster than basic Power Capping, it can still take up to one-half second to bring a server's power consumption back below the cap. As a result, it is possible to see peak power numbers reported that are several watts above a server's power cap value. Dynamic Power Capping still provides circuit protection, since it will always bring server power consumption below the cap long before a circuit breaker can trip.

Using HP Power Regulator in conjunction with power capping

Power capping and HP Power Regulator use the same underlying power management system to manage and control server power consumption. However, these two power management features are designed to accomplish different goals.

Power capping provides strict control of a server's maximum power consumption. It allows an administrator to set and maintain a limit on a server's maximum power consumption across a wider power range and without explicit regard to the impact on system performance. Power capping can use both processor P-states and clock throttling to limit system power consumption.

Power Regulator is focused on optimizing power use in a server (or group of servers) without impacting server performance. In its recommended configuration, Power Regulator dynamically selects the most power-efficient processor P-state that matches the present workload of the server. By doing this, Power Regulator lowers power consumption without affecting system performance or throughput. In this sense Power Regulator is primarily an efficiency tool.

Power Capping and HP Power Regulator can be used on an individual server or on a server group at the same time without creating a conflict. As long as a server's power consumption remains below the power cap setting, Power Regulator is free to manage the server using its settings and P-states. Power capping will override Power Regulator only if power consumption exceeds the established power cap.

Power capping and CPU utilization

Both power capping and Power Regulator involve manipulating the system CPU at a very basic level. As a result, power capping can affect CPU utilization numbers reported by monitoring tools at the operating system level. If a server under a relatively constant workload is assigned a power cap low enough to engage the power control mechanisms, monitoring utilities will generally report a greater CPU utilization number than they would if the cap were not in place.

Downloaded from www.Manualslib.com manuals search engine

Summary

HP Pawer Capping and HP Dynamic Pawer Capping are impartant pawer management features af HP PraLiant servers. Using pawer capping, system administratars can set limits an pawer cansumptian far a single server ar graups af servers that will nat affect server perfarmance while enabling them ta pravisian data center pawer and caaling resaurces ta apprapriate levels rather than the servers' maximum requirements. Enclasure Dynamic Pawer Capping represents the newest and mast saphisticated implementatian af pawer capping, with its ability ta actively reallacate an enclasure's pawer cap acrass individual server blades based an their warklaad requirements.

For more information

Far additianal infarmatian, refer ta the resaurces listed belaw.

Source	Hyperlink
HP Power Management	http://www.hp.com/go/powercapping
HP Insight Power Manager	http://www.hp.com/go/ipm
HP Integrated Lights-Out	http://www.hp.com/go/ilo
HP SIM page on hp.com	http://www.hp.com/go/hpsim
Dynamic Power Capping TCO and Best Practices white paper	http://h71028.www7.hp.com/ERC/downloads/4AA2-3107ENW.pdf
Dynamic Power Capping support matrix	http://h18004.www1.hp.com/products/servers/management/dynamic-power-capping/support.html
HP Power Calculators page	http://h30099.www3.hp.com/configurator/powercalcs.asp
HP ProLiant Energy Efficient Solutions	www.hp.com/go/proliant-energy-efficient
Power Regulator for ProLiant servers technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00300430/c00300430.pdf
Power Regulator	http://h18013.www1.hp.com/products/servers/management/ilo/power-regulator.html

Call to action

Send comments abaut this paper ta TechCam@HP.cam.

©2008, 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

AMD is a trademark of Advanced Micro Devices, Inc.

Intel is a trademark of Intel Corporation or its subsidiaries in the United States and other countries and is used under license.

TC090303TB, March 2009

Downloaded from www.Manualslib.com manuals search engine